

# Bertrand's postulate and subgroup growth

K. Bou-Rabee and D. B. McReynolds

September 7, 2009

## Abstract

In this article we investigate the  $L^1$ -norm of certain functions on groups called divisibility functions. Using these functions, their connection to residual finiteness, and integration theory on profinite groups, we define the residual average of a finitely generated group. One of the main results in this article is the finiteness of residual averages on finitely generated linear groups. Whether or not the residual average is finite depends on growth rates of indices of finite index subgroups. Our results on index growth rates are analogous to results on gaps between primes, and provide a variant of the subgroup growth function, which may be of independent interest.

keywords: *Bertrand's postulate, residual finiteness, subgroup growth*

MSC code: 20E07, 20E18

## 1 Introduction and main results

The study of gaps between successive primes has been a central topic in number theory for more than a hundred years. One classical result known as *Bertrand's postulate* asserts that for any positive integer  $n$ , there exists a prime between  $n$  and  $2n$ . This assertion was conjectured by Bertrand in 1845 and proved by Chebyshev in 1850 (see [26]). The stronger Legendre conjecture asserts a maximum gap of  $\sqrt{n}$  (see the related [15]), while the Prime Number Theorem provides many primes in the interval  $[n, 2n]$ .

One of the main purposes of this article is to investigate related problems for finitely generated groups. The role of primes in our setting will be played by relatively prime families of finite index subgroups  $\{\Delta_j\}$  or, more specifically,

the indices of such subgroups. Here, we say a family in  $\Gamma$  is *relatively prime* if for any distinct pair  $\Delta_j, \Delta_k$ , we have  $\Gamma = \Delta_j \Delta_k$ ; we will work with a stronger property called *prime* where a Chinese Remainder Theorem holds (see Section 3 for a precise definition). Our interest will be in proving results in the same vein as Bertrand's postulate.

**Question 1.** (Bertrand's postulate; classical) *On a finitely generated group  $\Gamma$ , when does there exist a prime (or normal) family of finite index subgroups  $\{\Delta_j\}$  and a constant  $d > 1$  such that*

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq d[\Gamma : \Delta_j]?$$

In the case  $\Gamma = \mathbf{Z}$  and the prime family of subgroups is  $\{p\mathbf{Z}\}$ , an affirmative answer to Question 1 is given by Bertrand's postulate. We could be less ambitious and allow ourselves small powers.

**Question 2.** (Bertrand's postulate; small powers) *On a finitely generated group  $\Gamma$ , when does there exist a prime (or normal) family of finite index subgroups  $\{\Delta_j\}$  and a constant  $\delta > 0$  such that*

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq [\Gamma : \Delta_j]^{1+\delta}?$$

Our first result resolves Question 1 for finitely generated linear groups.

**Theorem 1.1.** *Let  $\Gamma$  be an infinite finitely generated linear group over  $\mathbf{C}$ . Then there exists a constant  $d$  and an infinite family of finite index subgroups  $\{\Delta_j\}$  such that*

$$[\Gamma : \Delta_j] < [\Gamma : \Delta_{j+1}] \leq d[\Gamma : \Delta_j].$$

*In addition, there exists a finite index subgroup  $\Gamma_0$  of  $\Gamma$  such that  $\{\Delta_j\}$  is a normal, prime family in  $\Gamma_0$ .*

The proof of Theorem 1.1 uses the Lubotzky Alternative, the Strong Approximation Theorem, and Bertrand's postulate. We also prove the existence of families that resolve Question 2 that avoids the use of the Lubotzky Alternative and the Strong Approximation Theorem.

**Theorem 1.2.** *Let  $\Gamma$  be an infinite finitely generated linear group over  $\mathbf{C}$ . Then for any  $\delta > 0$ , there exists an infinite normal family of finite index subgroups  $\{\Delta_j\}$  such that*

$$[\Gamma : \Delta_j] < [\Gamma : \Delta_{j+1}] < [\Gamma : \Delta_j]^{1+\delta}.$$

It follows from [17] that in general finitely generated groups fail to have an affirmative answer to much weaker versions of Question 2. Indeed, for any increasing function  $f$ , there exists a finitely generated group  $\Gamma$  such that for any family of relatively prime subgroups  $\{\Delta_j\}$  there exist  $j_0$  such that

$$f([\Gamma : \Delta_{j_0}]) < [\Gamma : \Delta_{j_0+1}].$$

These results, aside from drawing analogies with the distribution of primes, are connected to a variant of subgroup growth that measures both relative index growth and the interplay between finite index subgroups. For this discussion we require some additional notation and terminology. Given a finitely generated, residually finite group  $\Gamma$ , let  $\text{FI}(\Gamma) = \{\Delta_j\}$  denote the set of finite index subgroups of  $\Gamma$  enumerated by index. Similarly,  $\text{NFI}(\Gamma)$  denotes the subcollection of normal, finite index subgroups. For a set  $S$  of  $\Gamma$ , we denote  $S \setminus \{1\}$  by  $S^\bullet$ . We define the *divisibility function*

$$D_\Gamma: \Gamma^\bullet \longrightarrow \mathbf{N}$$

by

$$D_\Gamma(\gamma) = \min \{ [\Gamma : \Delta_j] : \gamma \notin \Delta_j, \Delta_j \in \text{FI}(\Gamma) \}.$$

We define the associated *normal divisibility function* for normal, finite index subgroups in an identical way and denote it by  $D_\Gamma^\triangleleft$ .

In this article we study the  $L^1$ -norm of these and more general divisibility functions. Rivin [23] studied a similar norm on free groups, while the asymptotic behavior of  $L^\infty$ -norm on metric  $n$ -balls was the focus of the articles [3], [4], [13], and [23]. More explicitly, we define the *residual average*  $\text{Ave}(\Gamma)$  to be

$$\text{Ave}(\Gamma) = \int_{\widehat{\Gamma}} D_{\widehat{\Gamma}} d\mu,$$

where  $\widehat{\Gamma}$  is the profinite completion of  $\Gamma$ , the measure  $\mu$  is the Haar probability measure on  $\widehat{\Gamma}$ , and  $D_{\widehat{\Gamma}}$  is a continuous extension of  $D_\Gamma$  to  $\widehat{\Gamma}$ . Similarly, the *normal residual average* is

$$\text{Ave}^\triangleleft(\Gamma) = \int_{\widehat{\Gamma}} D_{\widehat{\Gamma}}^\triangleleft d\mu.$$

We will relate the finiteness of the above integrals to Questions 1 and 2—see Proposition 3.4. In particular, an application of Theorem 1.1 yields our next theorem.

**Theorem 1.3.** *If  $\Gamma$  is a finitely generated linear group over  $\mathbf{C}$ , then  $\text{Ave}(\Gamma)$  is finite.*

Using Theorem 1.2, we derive the following theorem.

**Theorem 1.4.** *If  $\Gamma$  is a finitely generated linear group over  $\mathbf{C}$ , then  $\text{Ave}^{\triangleleft}(\Gamma)$  is finite.*

The key ingredient in the proof of Theorem 1.3 is Theorem 1.1; likewise Theorem 1.2 is the key ingredient in the proof of Theorem 1.4. As the story leading to these connections is somewhat involved, we postpone a discussion of this here, hoping we have intrigued the reader sufficiently.

We mention one complementary result to Theorem 1.4. Namely, if  $\Gamma$  is either the first Grigorchuk group or  $\text{SL}(n, \mathbf{Z}_p)$  for  $n > 1$ , then  $\text{Ave}^{\triangleleft}(\Gamma)$  is infinite—see Theorem 7.1. These examples show that both linearity and finite generation are necessary in Theorem 1.4.

**Article layout.** In Section 2, we develop the framework for residual averages. In Section 3, we relate finiteness of residual averages to gaps between subgroups. In Section 4, we prove Theorems 1.1 and 1.3. In Section 5, we prove Theorem 1.2. In Section 6, we prove Theorem 1.4. In Section 7, we discuss the examples above. We discuss integrating over other densities like the asymptotic, annular, and spherical densities in Section 8. Finally, in Section 9 we briefly mention connections this article has with certain zeta functions studied by Larsen [18].

**Acknowledgements.** Foremost, we are extremely grateful to Benson Farb for his inspiration, comments, and guidance. We would also like to thank Emmanuel Breuillard for pointing out to us [5, Lemma 3.1]. In addition, we would like to thank Misha Belolipetsky, Emmanuel Breuillard, Jordan Ellenberg, Skip Garibaldi, Tsachik Gelander, Frédéric Haglund, Martin Kassabov, Jim Kelliher, Larsen Louder, Riad Masri, Alan Reid, and Tomasz Zamojski for numerous invaluable conversations on the material of this article. Finally, the second author was partially supported by an NSF postdoctoral fellowship.

## 2 Preliminaries

We begin with a section that constructs a rigorous framework for residual averages.

**1. Residual systems and divisibility functions.** Throughout,  $\Gamma$  will be an infinite, finitely generated, residually finite group. A collection  $\mathcal{F} = \{\Delta_j\}$  of finite index subgroups is a *residual system* if

$$\bigcap_{j=1}^{\infty} \Delta_j = \{1\}.$$

In addition, a residual system is *relative prime* if each distinct pair  $\Delta_j, \Delta_k$  in  $\mathcal{F}$  satisfies  $\Gamma = \Delta_j \Delta_k$ . It will be convenient to work with residual systems comprised of normal subgroups; we call such systems *normal residual systems*.

Associated to a residual system  $\mathcal{F}$  is the  $\mathcal{F}$ -divisibility function

$$D_{\mathcal{F}}(\gamma) = \min \{ [\Gamma : \Delta_j] : \gamma \notin \Delta_j, \Delta_j \in \mathcal{F} \}.$$

**2. Profinite completions.** Let  $\mathcal{T}_{\text{pro}}$  be the profinite topology on  $\Gamma$  given by declaring the finite index subgroups of  $\Gamma$  to be a neighborhood basis for the identity element and by declaring left multiplication to be a homeomorphism. This topology is also the weak topology on the set of all homomorphisms of  $\Gamma$  to finite groups, where we topologize the finite groups with the discrete topology  $\mathcal{T}_{\text{discrete}}$ .

There are several equivalent views of the profinite completion  $\widehat{\Gamma}$  of  $\Gamma$ . The profinite completion  $\widehat{\Gamma}$  is defined to be the inverse limit of the inverse limit system comprised of all finite quotients of  $\Gamma$ , where the finite quotients are equipped with the discrete topology. Consequently,  $\widehat{\Gamma}$  is a compact Hausdorff, topological group. We can also define the profinite completion to be the Cauchy completion of  $\Gamma$  with respect to either a uniform structure on  $\mathcal{T}_{\text{pro}}$  (see [14]) or via a metrization of  $\Gamma$  (either can be used to define equivalent notions of Cauchy for sequences). For the latter, the finite generation of  $\Gamma$  is required. We refer the reader to [29] for the general theory of profinite groups and profinite completions.

Set  $(\widehat{\Gamma}, \varphi)$  to be the profinite completion of  $\Gamma$  with associated continuous homomorphism

$$\varphi: \Gamma \longrightarrow \widehat{\Gamma}.$$

The image of  $\Gamma$  is dense and in the event that  $\Gamma$  is residually finite,  $\varphi$  is injective. By work of Haar [12], since  $\widehat{\Gamma}$  is a compact topological group,  $\widehat{\Gamma}$  is endowed with a bi-invariant Haar measure  $\mu$ . Upon normalizing, we assume throughout that  $\mu$  is the unique probability measure.

Associated to every finite index, normal subgroup  $\Delta$  in  $\text{NFI}(\Gamma)$  is an associated compact, open normal subgroup  $\overline{\Delta}$  in  $\widehat{\Gamma}$  defined by taking the closure of  $\varphi(\Delta)$  in  $\widehat{\Gamma}$ . The subgroup  $\overline{\Delta}$  yields an extension

$$\widehat{q}_\Delta: \widehat{\Gamma} \longrightarrow \Gamma/\Delta$$

of the canonical epimorphism

$$q_\Delta: \Gamma \longrightarrow \Gamma/\Delta$$

that satisfies  $q_\Delta = \widehat{q}_\Delta \circ \varphi$  (see [29, Proposition 1.4.2]).

We require the following elementary lemma for computational purposes.

**Lemma 2.1.** *Let  $\Delta_1, \Delta_2$ , be finite index subgroups of a residually finite group  $\Gamma$ . Then*

$$\overline{\Delta_1 \cap \Delta_2} = \overline{\Delta}_1 \cap \overline{\Delta}_2.$$

Lemma 2.1 follows from the well known fact that there exists a bijection between finite index, normal subgroups of  $\Gamma$  and compact, open, normal subgroups of  $\widehat{\Gamma}$ .

**3. Extending divisibility functions to profinite completions.** Having laid the groundwork for residual averages, we next extend  $D_\Gamma^\triangleleft$  to  $\widehat{\Gamma}$ . The first step is the following lemma.

**Lemma 2.2.**

$$D_\Gamma^\triangleleft: (\Gamma^\bullet, \mathcal{T}_{pro}) \longrightarrow (\mathbf{N}, \mathcal{T}_{discrete})$$

is continuous.

For the proof of the lemma, recall that  $\text{NFI}(\Gamma) = \{\Delta_j\}$  is the collection of normal, finite index subgroups ordered such that for all  $j$ , the inequality

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}]$$

holds. For each  $j$ , define

$$\Lambda_j = \bigcap_{\ell=1}^j \Delta_\ell, \quad \Gamma_j = \Lambda_{j-1} \Delta_j, \tag{1}$$

and

$$\text{NFI}_n(\Gamma) = \{\Delta \in \text{NFI}(\Gamma) : [\Gamma : \Delta] = n\}.$$

We now prove Lemmas 2.2.

*Proof.* We must show that for any subset  $S$  of  $\mathbf{N}$ , the pullback  $(D_\Gamma^\triangleleft)^{-1}(S)$  is open. As

$$(D_\Gamma^\triangleleft)^{-1}(S) = \bigcup_{s \in S} (D_\Gamma^\triangleleft)^{-1}(s),$$

it suffices to show that  $(D_\Gamma^\triangleleft)^{-1}(s)$  is open for all  $s$  in  $\mathbf{N}$ . To this end, note that

$$(D_\Gamma^\triangleleft)^{-1}(s) = \bigcup_{\Delta_j \in \text{NFI}_s(\Gamma)} (\Delta_j^c \cap \Lambda_{j-1}),$$

where  $\Delta_j^c$  denotes  $\Gamma \setminus \Delta_j$ . As  $\Delta_j$ ,  $\Lambda_j$  are both open and closed in the profinite topology,  $(D_\Gamma^\triangleleft)^{-1}(s)$  is open.  $\square$

Since any sequence  $\{\gamma_j\}$  in  $\Gamma^\bullet$  which converges to the identity in  $\mathcal{T}_{\text{pro}}$  has the property that  $\{D_\Gamma^\triangleleft(\gamma_j)\}$  converges to infinity, we continuously extend  $D_\Gamma^\triangleleft$  to  $\Gamma$  by declaring  $D_\Gamma^\triangleleft(1)$  to be infinity. The universal mapping property for profinite completions affords us with a unique, continuous extension

$$\widehat{D}_\Gamma^\triangleleft: \widehat{\Gamma} \longrightarrow \mathbf{N} \cup \{\infty\}$$

such that the diagram

$$\begin{array}{ccc} \widehat{\Gamma} & \xrightarrow{\widehat{D}_\Gamma^\triangleleft} & \mathbf{N} \cup \{\infty\} \\ \varphi \uparrow & \nearrow D_\Gamma^\triangleleft & \\ \Gamma & & \end{array}$$

commutes. This yields a continuous map

$$\widehat{D}_\Gamma^\triangleleft: \widehat{\Gamma} \longrightarrow \mathbf{N} \cup \{\infty\},$$

where the target is giving the topology induced by the 1-point compactification of  $\mathbf{R}$ . As a result, we know that  $\widehat{D}_\Gamma^\triangleleft$  is a measurable function and the integral

$$\int_{\widehat{\Gamma}} \widehat{D}_\Gamma^\triangleleft d\mu$$

is well defined. We define the *normal residual average* of  $\Gamma$  to be the integral

$$\text{Ave}^\triangleleft(\Gamma) = \int_{\widehat{\Gamma}} \widehat{D}_\Gamma^\triangleleft d\mu. \quad (2)$$

**Remark.** By uniqueness of the extension and the universal mapping property, we see that  $\widehat{D}_{\widehat{\Gamma}}^{\triangleleft} = D_{\widehat{\Gamma}}^{\triangleleft}$ , where

$$D_{\widehat{\Gamma}}^{\triangleleft}(\gamma) = \min \left\{ [\widehat{\Gamma} : \overline{\Delta_j}] : \overline{\Delta_j} \in \text{NFI}(\widehat{\Gamma}) \right\}.$$

For the remainder of the article, we will denote the extension by  $D_{\widehat{\Gamma}}^{\triangleleft}$ .

The next lemma provides a sum formula for  $\text{Ave}^{\triangleleft}(\Gamma)$ .

**Lemma 2.3.** *For  $\Gamma, \Delta_j$ , and  $\Lambda_j$  defined by (1), we have*

$$\int_{\widehat{\Gamma}} D_{\widehat{\Gamma}}^{\triangleleft} d\mu = \sum_{j=1}^{\infty} [\Gamma : \Delta_j] \left( \frac{[\Lambda_{j-1} : \Lambda_j] - 1}{[\Lambda_{j-1} : \Lambda_j]} \right) \left( \frac{1}{[\Gamma : \Lambda_{j-1}]} \right).$$

*Proof.* By the remark above, the definition of the Lebesgue integral, and properties of  $\widehat{\Gamma}$ , we see that

$$\begin{aligned} \int_{\widehat{\Gamma}} D_{\widehat{\Gamma}}^{\triangleleft} d\mu &= \sum_{n=1}^{\infty} n \mu((D_{\widehat{\Gamma}}^{\triangleleft})^{-1}(n)) \\ &= \sum_{n=1}^{\infty} \sum_{\Delta_j \in \text{NFI}_n(\Gamma)} [\widehat{\Gamma} : \overline{\Delta_j}] \mu(\overline{\Lambda_{j-1}} \setminus \overline{\Lambda_j}). \end{aligned}$$

An elementary calculation yields

$$\mu(\overline{\Lambda_{j-1}} \setminus \overline{\Lambda_j}) = \left( \frac{[\overline{\Lambda_{j-1}} : \overline{\Lambda_j}] - 1}{[\overline{\Lambda_{j-1}} : \overline{\Lambda_j}]} \right) \left( \frac{1}{[\widehat{\Gamma} : \overline{\Lambda_{j-1}}]} \right).$$

By Lemma 2.1, we get

$$\left( \frac{[\overline{\Lambda_{j-1}} : \overline{\Lambda_j}] - 1}{[\overline{\Lambda_{j-1}} : \overline{\Lambda_j}]} \right) \left( \frac{1}{[\widehat{\Gamma} : \overline{\Lambda_{j-1}}]} \right) = \left( \frac{[\Lambda_{j-1} : \Lambda_j] - 1}{[\Lambda_{j-1} : \Lambda_j]} \right) \left( \frac{1}{[\Gamma : \Lambda_{j-1}]} \right),$$

finishing the proof of the lemma.  $\square$

Replacing  $D_{\widehat{\Gamma}}^{\triangleleft}$  with  $D_{\widehat{\Gamma}}$  in the above discussion, yields a continuous extension  $D_{\widehat{\Gamma}}$  of  $D_{\widehat{\Gamma}}$ . We define the *residual average* of  $\Gamma$  to be

$$\text{Ave}(\Gamma) = \int_{\widehat{\Gamma}} D_{\widehat{\Gamma}} d\mu.$$

Lemma 2.3 also holds for  $D_{\widehat{\Gamma}}$ . Finally, since  $D_{\widehat{\Gamma}} \leq D_{\widehat{\Gamma}}^{\triangleleft}$ , note that

$$\text{Ave}(\Gamma) \leq \text{Ave}^{\triangleleft}(\Gamma).$$

**4. General residual systems.** For a normal, residual system  $\mathcal{F}$  on  $\Gamma$ , it need not be the case that the associated system  $\widehat{\mathcal{F}} = \{\widehat{\Delta}\}_{\Delta \in \mathcal{F}}$  is a residual system. In fact, if we take the completion  $\text{Cl}_{\mathcal{F}}(\Gamma)$  with respect to the system  $\mathcal{F}$ , by the universal mapping property for the profinite completion, we have a surjective homomorphism

$$\widehat{\varphi}_{\mathcal{F}}: \widehat{\Gamma} \longrightarrow \text{Cl}_{\mathcal{F}}(\Gamma),$$

and  $\ker \widehat{\varphi}_{\mathcal{F}}$  measures the failure of  $\widehat{\mathcal{F}}$  to be a residual system. On the other hand, we would like to define the  $\mathcal{F}$ -residual average of the  $\mathcal{F}$ -divisibility function  $D_{\mathcal{F}}$ . To this end, we define

$$\widehat{D}_{\mathcal{F}} = D_{\text{Cl}_{\mathcal{F}}(\Gamma)} \circ \widehat{\varphi}_{\mathcal{F}},$$

where  $D_{\text{Cl}_{\mathcal{F}}(\Gamma)}$  is the extension of  $D_{\mathcal{F}}$  to  $\text{Cl}_{\mathcal{F}}(\Gamma)$ . The construction of  $D_{\text{Cl}_{\mathcal{F}}(\Gamma)}$  is done in precisely the same way as the extension of  $D_{\Gamma}^{\triangleleft}$  to  $\widehat{\Gamma}$  was constructed above. We assert that

$$\int_{\widehat{\Gamma}} \widehat{D}_{\mathcal{F}} d\mu = \int_{\text{Cl}_{\mathcal{F}}(\Gamma)} D_{\text{Cl}_{\mathcal{F}}(\Gamma)} d\mu_{\mathcal{F}}, \quad (3)$$

where  $\mu_{\mathcal{F}}$  is the associated probability measure on  $\text{Cl}_{\mathcal{F}}(\Gamma)$ . To prove this equality, note that the map  $\widehat{\varphi}_{\mathcal{F}}$  has the property that for any measurable set  $A$  in  $\text{Cl}_{\mathcal{F}}(\Gamma)$ , we have the equality

$$\mu((\widehat{\varphi}_{\mathcal{F}})^{-1}(A)) = \mu_{\mathcal{F}}(A).$$

Consequently, for any measurable function  $g$  on  $\text{Cl}_{\mathcal{F}}(\Gamma)$ , we have the equality

$$\int_{\widehat{\Gamma}} g \circ \widehat{\varphi}_{\mathcal{F}} d\mu = \int_{\text{Cl}_{\mathcal{F}}(\Gamma)} g d\mu_{\mathcal{F}}.$$

Taking  $g$  to be  $D_{\text{Cl}_{\mathcal{F}}(\Gamma)}$  yields (3). In addition, the induced map

$$\widehat{\varphi}_{\mathcal{F}}^*: L^1(\text{Cl}_{\mathcal{F}}(\Gamma), \mu_{\mathcal{F}}) \longrightarrow L^1(\widehat{\Gamma}, \mu)$$

is an isometric embedding, where for  $g \in L^1(\text{Cl}_{\mathcal{F}}(\Gamma), \mu_{\mathcal{F}})$ , we have

$$\widehat{\varphi}_{\mathcal{F}}^*(g) = g \circ \widehat{\varphi}_{\mathcal{F}}.$$

We define the  $\mathcal{F}$ -residual average to be

$$\text{Ave}_{\mathcal{F}}(\Gamma) = \int_{\widehat{\Gamma}} \widehat{D}_{\mathcal{F}} d\mu.$$

As before, we have

$$\text{Ave}_{\mathcal{F}}(\Gamma) = \sum_{\Delta_j \in \mathcal{F}} [\Gamma : \Delta_j] \left( \frac{[\Lambda_{j-1} : \Lambda_j] - 1}{[\Lambda_{j-1} : \Lambda_j]} \right) \left( \frac{1}{[\Gamma : \Lambda_{j-1}]} \right),$$

where  $\Lambda_j$  is the intersection over the first  $j$  subgroups in  $\mathcal{F}$ .

**Lemma 2.4.** *If  $\mathcal{F}$  is any residual system on  $\Gamma$  for which  $\text{Ave}_{\mathcal{F}}(\Gamma)$  is finite, then  $\text{Ave}(\Gamma)$  is finite.*

**Lemma 2.5.** *Let  $\psi: \Gamma \rightarrow \Lambda$  be a surjective homomorphism of finitely generated, residually finite groups. If  $\Lambda$  is infinite and  $\text{Ave}(\Lambda)$  is finite, then  $\text{Ave}(\Gamma)$  is finite. In addition, if  $\text{Ave}^{\triangleleft}(\Lambda)$  is finite, then  $\text{Ave}^{\triangleleft}(\Gamma)$  is finite.*

As the validity of Lemma 2.4 is clear, we only prove Lemma 2.5.

*Proof of Lemma 2.5.* To begin, by the universal mapping property for profinite completions, the homomorphism  $\psi$  has a continuous extension

$$\widehat{\psi}: \widehat{\Gamma} \longrightarrow \widehat{\Lambda}.$$

The map  $\widehat{\psi}$  induces an isometric embedding

$$\widehat{\psi}^*: L^1(\widehat{\Lambda}, \mu_{\widehat{\Lambda}}) \longrightarrow L^1(\widehat{\Gamma}, \mu_{\widehat{\Gamma}}).$$

In particular, by hypothesis  $\widehat{\psi}^*(D_{\widehat{\Lambda}})$ ,  $\widehat{\psi}^*(D_{\widehat{\Lambda}}^{\triangleleft})$  are in  $L^1(\widehat{\Gamma}, \mu_{\widehat{\Gamma}})$ . The proof is completed by noting the inequalities

$$D_{\widehat{\Gamma}} \leq \widehat{\psi}^*(D_{\widehat{\Lambda}}) \text{ and } D_{\widehat{\Gamma}}^{\triangleleft} \leq \widehat{\psi}^*(D_{\widehat{\Lambda}}^{\triangleleft}).$$

□

### 3 Gaps between subgroups and residual averages

In this section, we relate the finiteness of residual averages to gaps between subgroups. This connection will be done via elementary group theory and Lemma 2.3.

**1. Index sum formula.** We begin by deriving a formula for  $\text{Ave}_{\mathcal{F}}(\Gamma)$  in terms of indices of subgroups when  $\mathcal{F} = \{\Delta_j\}$  is a normal residual system. For  $\Delta_j$ , we have (see (1))

$$\begin{array}{c}
 \Gamma \\
 \downarrow r_j \\
 \Gamma_j \\
 \swarrow t_j \quad \searrow s_j \\
 \Lambda_{j-1} \quad \Delta_j \\
 \searrow s_j \quad \swarrow t_j \\
 \Lambda_j
 \end{array} \tag{4}$$

and from this we obtain

$$\mu(\overline{\Lambda_{j-1}} \setminus \overline{\Lambda_j}) = \left( \frac{[\Lambda_{j-1} : \Lambda_j] - 1}{[\Lambda_{j-1} : \Lambda_j]} \right) \left( \frac{1}{[\Gamma : \Lambda_{j-1}]} \right) = \frac{s_j - 1}{r_j s_j t_j}.$$

As  $[\Gamma : \Delta_j] = r_j s_j$ , we obtain our next lemma.

**Lemma 3.1.**

$$\text{Ave}_{\mathcal{F}}(\Gamma) = \sum_{j=1}^{\infty} \frac{(s_j - 1)}{t_j}.$$

We say that a residual system  $\mathcal{F}$  is *nested* if  $\Delta_{j+1} < \Delta_j$  for all  $j$ . For a nested, normal residual system, from (4) it follows that  $t_j = 1$  for all  $j$ . In particular, by Lemma 3.1, we have the following lemma.

**Lemma 3.2.** *If  $\mathcal{F}$  is a nested, normal residual system on  $\Gamma$ , then  $\text{Ave}_{\mathcal{F}}(\Gamma)$  is infinite.*

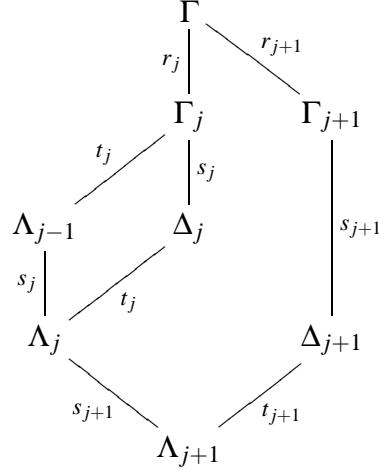
We will use this in Section 7 to show the first Grigorchuk group and  $\text{SL}(n, \mathbf{Z}_p)$  have infinite normal residual averages.

**2. Another index sum formula.** We next aim to recursively compute the coefficients  $t_j$ . This is achieved with the following.

**Lemma 3.3.** *Let  $\mathcal{F} = \{\Delta_j\}$  be a normal residual system. For all  $j$ , we have*

$$t_{j+1} = \frac{\prod_{\ell=1}^j s_{\ell}}{r_{j+1}}.$$

*Proof.* For this, consider the diagram:



As we have two paths from  $\Lambda_{j+1}$  to  $\Gamma$ , we see that

$$[\Gamma : \Lambda_{j+1}] = t_{j+1}s_{j+1}r_{j+1} = s_{j+1}t_js_jr_j,$$

and thus

$$t_{j+1} = \frac{r_js_jt_j}{r_{j+1}}. \quad (5)$$

To verify the formula for  $t_{j+1}$ , we utilize (5) via an inductive argument. For the base case, by definition,  $r_1 = 1$  and  $t_1 = 1$ , and so

$$t_2 = \frac{r_1s_1t_1}{r_2} = \frac{s_1}{r_2}.$$

We next assume now that the formula holds for  $j$ . From this assumption, we deduce the asserted formula:

$$t_{j+1} = \frac{r_js_jt_j}{r_{j+1}} = \left( \frac{r_js_j}{r_{j+1}} \right) \left( \frac{\prod_{\ell=1}^{j-1} s_\ell}{r_j} \right) = \frac{\prod_{\ell=1}^j s_\ell}{r_{j+1}}.$$

□

From Lemmas 3.1 and 3.3, we obtain another index sum formula for  $\text{Ave}_{\mathcal{F}}(\Gamma)$ :

$$\text{Ave}_{\mathcal{F}}(\Gamma) = \sum_{j=1}^{\infty} \frac{r_j(s_j-1)}{\prod_{\ell=1}^{j-1} s_\ell}. \quad (6)$$

**3. Index growth and convergence.** We are now in position to relate the finiteness of  $\text{Ave}(\Gamma)$  to gaps in subgroup growth. With  $r_j, s_j$  given by (4), define the sequence

$$\alpha_j(\mathcal{F}) = \frac{r_{j+1}(s_{j+1} - 1)}{r_j s_j(s_j - 1)}.$$

We say that  $\mathcal{F}$  as *sub-quadratic index growth* if  $\alpha_j(\mathcal{F}) < 1$  for sufficiently large  $j$ . We say that  $\mathcal{F}$  as *super-quadratic index growth* if  $\alpha_j(\mathcal{F}) > 1$  for sufficiently large  $j$ .

**Proposition 3.4.** *Let  $\Gamma$  be a finitely generated, residually finite group and  $\mathcal{F}$  a normal residual system.*

- (a) *If  $\mathcal{F}$  has sub-quadratic index growth, then  $\text{Ave}_{\mathcal{F}}(\Gamma)$  is finite.*
- (b) *If  $\mathcal{F}$  has super-quadratic index growth, then  $\text{Ave}_{\mathcal{F}}(\Gamma)$  is infinite.*

*Proof.* According to (6),

$$\text{Ave}_{\mathcal{F}}(\Gamma) = \sum_{j=1}^{\infty} \frac{r_j(s_j - 1)}{\prod_{\ell=1}^{j-1} s_{\ell}},$$

where  $r_j, s_j$  are given in (4). We see that the ratio of consecutive terms is given by

$$\frac{r_{j+1}(s_{j+1} - 1)}{\prod_{\ell=1}^j s_{\ell}} \left( \frac{r_j(s_j - 1)}{\prod_{\ell=1}^{j-1} s_{\ell}} \right)^{-1} = \frac{r_{j+1}(s_{j+1} - 1)}{r_j s_j(s_j - 1)},$$

and so both (a) and (b) follow from the Ratio Test.  $\square$

We call a normal residual system  $\mathcal{F}$  *prime* if  $r_j = 1$  for all  $j$ . One reason for this terminology is that prime systems have a property analogous to the Chinese Remainder Theorem:

$$\Gamma/\Lambda_j = \bigoplus_{\ell=1}^j \Gamma/\Delta_{\ell}. \quad (7)$$

For such residual systems, the subgroups  $\Delta_j$  have minimal overlap. Moreover, as  $[\Gamma : \Delta_j]$  is unbounded, Proposition 3.4 reduces to studying the limit of the sequence

$$\alpha_j(\mathcal{F}) = \frac{[\Gamma : \Delta_{j+1}]}{[\Gamma : \Delta_j]^2}.$$

Therefore, the finiteness of  $\text{Ave}_{\mathcal{F}}(\Gamma)$  depends only on the growth between consecutive indices for prime systems.

For a normal residual system the coefficient  $r_j$  measures the failure of (7) at the  $j$ th stage. Consequently, residual averages depend on the intersection of pairs and not just the growth of indices in general.

For a non-normal residual system  $\mathcal{F} = \{\Delta_j\}$ , we say that  $\mathcal{F}$  is *prime* if for all  $j$ , we have

$$[\Gamma : \Lambda_j] = \prod_{\ell=1}^j [\Gamma : \Delta_\ell].$$

One case when this holds is when the indices  $[\Gamma : \Delta_j]$  are pairwise relatively prime.

**4. Residual averages on  $\mathbf{Z}$ .** For  $\Gamma = \mathbf{Z}$ , divisibility functions have simple interpretations. For an integer  $m$ , we define three functions

$$\begin{aligned} D(m) &= \min \{n : \gcd(m, n) < \min \{m, n\}, n > 1\} \\ D_{\text{prime}}(m) &= \min \{p : \gcd(m, p) = 1, p \text{ prime}\} \\ D_p(m) &= \min \{p^j : \gcd(m, p^j) = 1, p \text{ a fixed prime}\} \end{aligned}$$

We define the averages of these functions as before and denote them  $\text{Ave}(\mathbf{Z})$ ,  $\text{Ave}_{\text{prime}}(\mathbf{Z})$ , and  $\text{Ave}_p(\mathbf{Z})$ , respectively. In the statement of our next result,  $\{p_j\}$  denotes the set of primes ordered by cardinality.

**Proposition 3.5.** *We have the following formulas:*

$$\begin{aligned} \text{Ave}(\mathbf{Z}) &= \sum_{j=1}^{\infty} j \left( 1 - \frac{\text{lcm}(1, \dots, j-1)}{\text{lcm}(1, \dots, j)} \right) \left( \frac{1}{\text{lcm}(1, \dots, j-1)} \right) \\ \text{Ave}_{\text{prime}}(\mathbf{Z}) &= \sum_{j=1}^{\infty} \frac{p_j - 1}{\prod_{\ell=1}^{j-1} p_\ell}, \quad \text{Ave}_p(\mathbf{Z}) = \sum_{j=1}^{\infty} (p-1) \end{aligned}$$

In particular,  $\text{Ave}(\mathbf{Z})$ ,  $\text{Ave}_{\text{prime}}(\mathbf{Z})$  are finite while  $\text{Ave}_p(\mathbf{Z})$  is not finite for any prime  $p$ .

This proposition follows easily using Bertrand's postulate, the Prime Number Theorem, and Proposition 3.4. We also have (see Rivin [23] for the second series value):

$$\text{Ave}(\mathbf{Z}) \approx 2.787780456, \quad \text{Ave}_{\text{prime}}(\mathbf{Z}) \approx 2.920050977.$$

## 4 Controlling gaps for linear groups

In this section, we prove Theorem 1.1. Our proof splits into two cases depending on whether or not  $\Gamma$  is virtually solvable.

**1. Virtually solvable groups.** When  $\Gamma$  is virtually solvable, Theorem 1.1 can be shown easily in a few different ways.

**Proposition 4.1.** *Let  $\Gamma$  be a virtually solvable, finitely generated linear group over  $\mathbb{C}$ . Then there exists a constant  $d$  and a family of finite index subgroups  $\{\Delta_j\}$  of  $\Gamma$  such that*

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq d[\Gamma : \Delta_j].$$

*In addition, there exists a finite index subgroup  $\Gamma_0$  of  $\Gamma$  such that  $\{\Delta_j\}$  is a normal, prime family in  $\Gamma_0$ .*

*Proof.* It is well known that there exists a finite index subgroup  $\Gamma_0$  of  $\Gamma$  with the property that there exists a surjective homomorphism

$$\varphi: \Gamma_0 \longrightarrow \mathbf{Z}.$$

Set  $\Delta_j = \ker r_{p_j} \circ \varphi$ , where  $\{p_j\}$  is the set of primes and

$$r_{p_j}: \mathbf{Z} \longrightarrow \mathbf{F}_{p_j}$$

is reduction modulo  $p_j$ . This is a normal, prime family and by Bertrand's postulate, we have

$$[\Gamma_0 : \Delta_j] \leq [\Gamma_0 : \Delta_{j+1}] \leq 2[\Gamma_0 : \Delta_j].$$

Viewed as subgroups of  $\Gamma$ , we obtain a family of finite index subgroups. Moreover,

$$\begin{aligned} [\Gamma_0 : \Delta_j] &\leq [\Gamma_0 : \Delta_{j+1}] \leq 2[\Gamma_0 : \Delta_j] \\ [\Gamma : \Gamma_0][\Gamma_0 : \Delta_j] &\leq [\Gamma : \Gamma_0][\Gamma_0 : \Delta_{j+1}] \leq 2[\Gamma : \Gamma_0][\Gamma_0 : \Delta_j] \\ [\Gamma : \Delta_j] &\leq [\Gamma : \Delta_{j+1}] \leq 2[\Gamma : \Delta_j], \end{aligned}$$

as needed.  $\square$

An alternative to the above proof is to realize a finite index subgroup  $\Gamma_0$  of  $\Gamma$  as a finite index subgroup of  $\mathbf{S}(\mathcal{O}_k)$  for a solvable, linear  $k$ -algebraic group  $\mathbf{S}$ . We

then use reduction homomorphism on  $\mathbf{S}(\mathcal{O}_k)$  to produce a prime, residual system with the desired index gaps on  $\Gamma_0$ . Hence, in Proposition 4.1, we also have

$$\bigcap_{j=1}^{\infty} \Delta_j = 1.$$

Finally, we can also arrange it so that the subgroups  $\Delta_j$  are normal in  $\Gamma$  by appealing to the congruence subgroup property for  $\mathbf{S}(\mathcal{O}_k)$  (see [8]). As neither of these properties are required in the sequel, we have opted to omit the details for these upgrades.

**2. A simple example.** For non-virtually solvable groups, we focus first on the case when  $\Gamma = \mathbf{G}(\mathbf{Z})$  for a connected, simply connected, simple, linear  $\mathbf{Q}$ -algebraic group  $\mathbf{G}$ . Before undertaking this endeavor, we present a simple, motivational example.

**Example.** Let  $\mathbf{G} = \mathrm{SL}(n, \mathbf{C})$  and  $\Gamma = \mathrm{SL}(n, \mathbf{Z})$ . For each prime  $p_j$ , we have a surjective homomorphism

$$r_{p_j} : \mathrm{SL}(n, \mathbf{Z}) \longrightarrow \mathrm{SL}(n, \mathbf{F}_{p_j})$$

given by reducing coefficients modulo  $p_j$ . A simple computation shows that

$$|\mathrm{SL}(n, \mathbf{F}_{p_j})| = \frac{\prod_{\ell=0}^{n-1} (p_j^n - p_j^\ell)}{p - 1} = [\Gamma : \ker r_{p_j}].$$

According to Bertrand's postulate, we know that

$$\frac{[\Gamma : \ker r_{p_{j+1}}]}{[\Gamma : \ker r_{p_j}]} \leq \frac{(p_j - 1) \prod_{\ell=0}^{n-1} ((2p_j)^n - (2p_j)^\ell)}{(2p_j - 1) \prod_{\ell=0}^{n-1} (p_j^n - p_j^\ell)}.$$

We have a prime normal residual system  $\{\Delta_j\}$  by setting  $\Delta_j = \ker r_{p_j}$ . Applying L'Hôpital's rule, we obtain

$$\lim_{j \rightarrow \infty} \frac{[\Gamma : \Delta_{j+1}]}{[\Gamma : \Delta_j]} \leq 2^{n^2 - 1}.$$

Therefore, for large values of  $j$ , we have

$$[\Gamma : \ker r_{p_j}] \leq [\Gamma : \ker r_{p_{j+1}}] \leq 2^{n^2} [\Gamma : \ker r_{p_j}] = 2^{\dim \mathbf{G} + 1} [\Gamma : \ker r_{p_j}].$$

**3. Integral points in simple linear algebraic groups.** Using the method above, we prove the following proposition.

**Proposition 4.2.** *Let  $\mathbf{G}$  be a connected, simply connected, simple, linear  $\mathbf{Q}$ -algebraic group and  $\Gamma = \mathbf{G}(\mathbf{Z})$ . Then there exists a constant  $d$  and a family of normal, prime, finite index subgroups  $\{\Delta_j\}$  of  $\Gamma$  such that*

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq d[\Gamma : \Delta_j].$$

In addition,

$$\bigcap_{j=1}^{\infty} \Delta_j = 1.$$

In the proof, we write  $f(p) \sim g(p)$  to mean  $\lim_{p \rightarrow \infty} \frac{f(p)}{g(p)} = 1$ , where the limit is taken over  $p$  in the intersection of the domains of  $f$  and  $g$ .

*Proof.* According to the Strong Approximation Theorem (see [21], [22], or [28]), for all but finitely many primes  $p$ , the reduction modulo  $p$  homomorphism maps  $\Gamma$  surjectively onto  $\mathbf{G}(\mathbf{F}_p)$ . In addition, the kernels  $\ker r_{p_j}$  yield a normal, prime, residual system on  $\Gamma$ . As  $|\mathbf{G}(\mathbf{F}_p)| \sim p^{\dim \mathbf{G}}$  (see [19, p. 123] or [25, p. 131]), by Bertrand's postulate, we see for consecutive primes  $p_j, p_{j+1}$  that

$$\begin{aligned} |\mathbf{G}(\mathbf{F}_{p_{j+1}})| &\sim p_{j+1}^{\dim \mathbf{G}} \\ &\leq (2p_j)^{\dim \mathbf{G}} \\ &= 2^{\dim \mathbf{G}} p_j^{\dim \mathbf{G}} \\ &\sim 2^{\dim \mathbf{G}} |\mathbf{G}(\mathbf{F}_{p_j})|. \end{aligned}$$

Thus, setting  $d = 2^{\dim \mathbf{G}+1}$ , we get the desired gap condition for the normal, prime family  $\{\ker r_{p_j}\}$  for sufficiently large  $j$ .  $\square$

**4. Non-solvable groups.** Using Proposition 4.2 and the Lubotzky Alternative, we obtain the following corollary.

**Corollary 4.3.** *Let  $\Gamma$  be a finitely generated linear group over  $\mathbf{C}$  that is not virtually solvable. Then there exists a constant  $d$  and a family of finite index subgroups  $\{\Delta_j\}$  of  $\Gamma$  such that*

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq d[\Gamma : \Delta_j].$$

*In addition, there exists a finite index subgroup  $\Gamma_0$  of  $\Gamma$  such that  $\{\Delta_j\}$  is a normal, prime family in  $\Gamma_0$ .*

*Proof.* As  $\Gamma$  is not virtually solvable, the Lubotzky Alternative (see [19, Theorem 16.4.12]) yields a finite index subgroup  $\Gamma_0$  of  $\Gamma$  and a representation

$$\rho : \Gamma_0 \longrightarrow \mathbf{G}(\mathbf{Z}_{\mathcal{P}_{\text{good}}}),$$

where  $\mathbf{G}$  is a connected, simply connected, simple, linear  $\mathbf{Q}$ -algebraic group,  $\mathcal{P}_{\text{good}}$  is a finite set of integral primes, and

$$\mathbf{Z}_{\mathcal{P}_{\text{good}}} = \bigoplus_{p \notin \mathcal{P}_{\text{good}}} \mathbf{Z}_p.$$

The content of the Lubotzky Alternative is that it guarantees that the Strong Approximation Theorem can be applied to  $\rho(\Gamma_0)$ . In particular, for all but finitely many primes (possibly more than  $\mathcal{P}_{\text{good}}$ ), reduction modulo  $p$  maps  $\Gamma_0$  onto  $\mathbf{G}(\mathbf{F}_p)$ . Taking  $\mathcal{F} = \{\ker r_{p_j} \circ \rho\}$ , we obtain a normal prime family on  $\Gamma_0$ . Moreover, by Proposition 4.2, we have

$$[\Gamma_0 : \Delta_j] \leq [\Gamma_0 : \Delta_{j+1}] \leq d[\Gamma_0 : \Delta_j].$$

Viewing these subgroups inside of  $\Gamma$ , as before, we see that

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq d[\Gamma : \Delta_j]$$

still holds. Thus, we have a family of finite index subgroups on  $\Gamma$  with the desired gap condition.  $\square$

It could very well be that the subgroups  $\Delta_j$  are not normal in  $\Gamma$ . Set

$$\text{Core}(\Delta_j) = \bigcap_{\gamma \in \Gamma} \gamma^{-1} \Delta_j \gamma$$

to be the normal core of  $\Delta_j$  in  $\Gamma$ . These subgroups yield the normal family  $\{\text{Core}(\Delta_j)\}$  on  $\Gamma$ . We cannot ensure the gap condition for the indices since

$$[\Gamma : \Delta_j] \leq [\Gamma : \text{Core}(\Delta_j)] \leq [\Gamma : \Delta_j]^{[\Gamma : \Gamma_0]}.$$

With regard to finiteness of normal residual averages, this is a problem.

**5. The proof of Theorem 1.1.** Let  $\Gamma$  be a finitely generated linear group over  $\mathbf{C}$ . Recall that for Theorem 1.1, we must produce a family of finite index subgroups  $\{\Delta_j\}$  such that

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq d[\Gamma : \Delta_j]$$

for some constant  $d$  and all  $j$ . If  $\Gamma$  is virtually solvable, the existence is the content of Proposition 4.1. Otherwise, the existence of such a family follows from Corollary 4.3.  $\square$

**6. Proof of Theorem 1.3.** According to Theorem 1.1, there exists a normal, finite index subgroup  $\Gamma_0$  of  $\Gamma$  and a normal, prime family of finite index subgroups  $\{\Delta_j\}$  on  $\Gamma_0$  such that

$$[\Gamma_0 : \Delta_j] \leq [\Gamma_0 : \Delta_{j+1}] \leq d[\Gamma_0 : \Delta_j].$$

Setting

$$K = \bigcap_{j=1}^{\infty} \Delta_j,$$

and  $\Lambda_0 = \Gamma_0/K$ , the family  $\{\Delta_j\}$  descends to a normal, prime residual system  $\{\Delta'_j\}$  on  $\Lambda_0$ . Since

$$\frac{[\Lambda_0 : \Delta'_{j+1}]}{[\Lambda_0 : \Delta'_j]^2} \leq \frac{d}{[\Lambda_0 : \Delta'_j]}$$

and  $[\Lambda_0 : \Delta'_j] > d$  for all but finitely many  $j$ , by Proposition 3.4 and Lemma 2.4,  $\text{Ave}(\Lambda_0)$  is finite. By construction,  $K$  has infinite index, and thus by Lemma 2.5,  $\text{Ave}(\Gamma_0)$  is finite. We claim now that this implies that  $\text{Ave}(\Gamma)$  is finite. For this, we have

$$\text{Ave}(\Gamma) = \int_{\widehat{\Gamma}} D_{\widehat{\Gamma}} d\mu = \int_{\widehat{\Gamma} \setminus \overline{\Gamma_0}} D_{\widehat{\Gamma}} d\mu + \int_{\overline{\Gamma_0}} D_{\widehat{\Gamma}} d\mu. \quad (8)$$

The finiteness of integrals on the right hand side of (8) can now be seen from the following two facts:

$$\int_{\widehat{\Gamma} \setminus \overline{\Gamma_0}} D_{\widehat{\Gamma}} d\mu \leq [\Gamma : \Gamma_0] \mu(\widehat{\Gamma} \setminus \overline{\Gamma_0}) < \infty,$$

and

$$\int_{\overline{\Gamma_0}} D_{\widehat{\Gamma}} d\mu \leq [\Gamma : \Gamma_0] \int_{\overline{\Gamma_0}} D_{\overline{\Gamma_0}} d\mu = [\Gamma : \Gamma_0] \text{Ave}(\Gamma_0) < \infty.$$

□

## 5 Proof of Theorem 1.2

In this section we prove Theorem 1.2.

**1. The main proposition.** The following proposition is the main step in proving Theorem 1.2.

**Proposition 5.1.** *Let  $\Gamma$  be a finitely generated subgroup of  $\mathrm{GL}(n, K)$  and  $K/\mathbf{Q}$  a finite extension. Then for each  $\delta > 0$ , there exists a normal residual system  $\mathcal{F}_\delta$  such that for all  $j$ ,*

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq [\Gamma : \Delta_j]^{1+\delta}.$$

As the proof is somewhat involved, we summarize our strategy for the reader. Using the group  $\mathrm{SL}(n, \mathbf{Z})$  as a model, we take kernels of reduction homomorphisms

$$r_{\mathfrak{p}}: \Gamma \longrightarrow \mathrm{GL}(n, S/\mathfrak{p})$$

for a particular subring  $S$  in  $K$  and prime ideals  $\mathfrak{p}$  of  $S$ . Unlike the case of  $\mathrm{SL}(n, \mathbf{Z})$ , we have no control here on the size of the index of  $\ker r_{\mathfrak{p}}$ . We circumvent this by instead taking reductions modulo  $\mathfrak{p}_j^{k_j}$  for suitable powers  $k_j$ . The selection of these powers comprises the bulk of the proof. With regard to exposition, the difficulty is the interdependence of several quantities, each of which requires control for the selection of the powers  $k_j$ . With this in mind, in the proof below, we indicate the dependence of  $x$  on  $y$  by  $x_y$ . We hope this makes clear the dependence of each quantity on the others. These dependencies are important in both the proof of Proposition 5.1 and of Theorem 1.4.

*Proof of Proposition 5.1.* For a finitely generated subgroup  $\Gamma$  of  $\mathrm{GL}(n, K)$  and a finite generating set  $\{\gamma_m\}$  of  $\Gamma$ , we define  $S$  to be the ring generated by  $\{(\gamma_m)_{i,j}\}$ . By possibly enlarging  $S$ , we can assume that  $\mathcal{O}_K$  is contained in  $S$ , where  $\mathcal{O}_K$  is the ring of  $K$ -integers. As  $\Gamma$  is finitely generated, the set of prime ideals of  $S$  can be identified with a co-finite subset of  $\mathcal{P}_K$ , where  $\mathcal{P}_K$  is the set of prime ideals of  $\mathcal{O}_K$ . Specifically, we associate to  $\mathfrak{p}$  in  $\mathcal{P}_K$  the ideal  $S\mathfrak{p} = \mathfrak{p}_S$ . For all but finitely many prime ideals  $\mathfrak{p}$  in  $\mathcal{O}_K$ , the ideal  $\mathfrak{p}_S$  is a proper, prime ideal and

$$S/\mathfrak{p}_S \cong \mathcal{O}_K/\mathfrak{p} \cong \mathbf{F}_{p^{\mathfrak{e}_{\mathfrak{p}}}}.$$

We denote the set of all prime ideals in  $S$  by  $\mathcal{P}_S$ . Excluding finitely many primes  $p$ , we select for each prime  $p$  in  $\mathbf{Z}$  a prime ideal  $\mathfrak{p}_S$  in  $S$  such that  $S/\mathfrak{p}_S$  has characteristic  $p$  and  $|S/\mathfrak{p}_S|$  is minimal. We denote this positive density subset of  $\mathcal{P}_S$  by  $\mathcal{P}_S^1$ . According to the Cebotarev Density Theorem, we can pass to a positive density subset  $\mathcal{P}_S^2$  of  $\mathcal{P}_S^1$  such that for all  $\mathfrak{p}_S$  in  $\mathcal{P}_S^2$ , we have

$$S/\mathfrak{p}_S = \mathbf{F}_p.$$

We order  $\mathcal{P}_S^2 = \{\mathfrak{p}_{S,j}\}$  via the characteristic of the associated residue fields  $\mathbf{F}_p$ . By the Prime Number Theorem (see the proof of Lemma 2.4 in [3]), the positive density of  $\mathcal{P}_S^2$  implies that there exists an integer  $d$  such that

$$p_j \leq p_{j+1} \leq dp_j \quad (9)$$

for all but finitely many  $j$ , where  $p_j = |S/\mathfrak{p}_{S,j}|$ . Enlarging  $d$  if necessary, we may assume that (9) holds for all  $j$ . For each  $\mathfrak{p}_{S,j}$  in  $\mathcal{P}_S^2$ , we have the reduction homomorphism

$$r_j: \mathrm{GL}(n, S) \longrightarrow \mathrm{GL}(n, S/\mathfrak{p}_{S,j}) = \mathrm{GL}(n, \mathbf{F}_{p_j})$$

given by reducing the coefficients modulo the prime  $\mathfrak{p}_{S,j}$ . By our selection of  $S$ , the group  $\Gamma$  is a subgroup of  $\mathrm{GL}(n, S)$ , and so the homomorphisms  $r_j$  endow  $\Gamma$  with a residual system  $\mathcal{F} = \{\ker r_j \cap \Gamma\}$ . To see that this is a residual system, notice that for any  $\gamma$  in  $\Gamma^\bullet$ , there are only finitely many prime ideals  $\mathfrak{p}_{S,j}$  such that  $\gamma$  resides in  $\ker r_j$ . In particular, we see that there are only finitely many  $j$  such that  $\Gamma$  is contained in  $\ker r_j$ , and we set  $\mathcal{P}_S^3$  to be the set of primes in  $\mathcal{P}_S^2$  for which  $\Gamma$  is not a subgroup of  $\ker r_j$ . Set  $\mathcal{F}' = \{\ker r_j \cap \Gamma\}$  for  $\mathfrak{p}_{S,j}$  in  $\mathcal{P}_S^3$ . By construction, the image of  $\Gamma$  under  $r_j$  is a nontrivial subgroup of  $\mathrm{GL}(n, \mathbf{F}_{p_j})$  and thus has order  $O_j$  where

$$1 < O_j \leq \prod_{\ell=0}^{n-1} (p_j^n - p_j^\ell) < p_j^{n^2}. \quad (10)$$

To avoid controlling each  $O_j$ , we instead pass to deeper subgroups of  $\Gamma \cap \ker r_j$  given by reduction modulo prime powers. If we reduce modulo the prime power  $\mathfrak{p}_{S,j}^{k_j}$ , we obtain the homomorphism

$$r_{j,k_j}: \mathrm{GL}(n, S) \longrightarrow \mathrm{GL}(n, S/\mathfrak{p}_{S,j}^{k_j}).$$

The image of  $\Gamma$  under  $r_{j,k_j}$  has order

$$|r_{j,k_j}(\Gamma)| = O_j p_j^{\ell_{j,k_j}}$$

This equality follows from the fact that for all  $k > 1$ , we have (see [1, Corollary 9.3], [10, Ch. 9], or the proof of Lemma 16.4.5 in [19])

$$1 \longrightarrow \mathfrak{gl}_n(\mathbf{F}_{p_j}) \longrightarrow \mathrm{GL}(n, S/\mathfrak{p}_{S,j}^k) \longrightarrow \mathrm{GL}(n, S/\mathfrak{p}_{S,j}^{k-1}) \longrightarrow 1,$$

where  $\mathfrak{gl}_n(\mathbf{F}_{p_j})$  is the Lie algebra  $\mathfrak{gl}_n$  of  $\mathrm{GL}_n$  with coefficients in the finite field  $\mathbf{F}_{p_j}$ . In particular, via induction, we have

$$\left| \mathrm{GL}(n, S/\mathfrak{p}_{S,j}^{k_j}) \right| = p_j^{n^2(k_j-1)} \left| \mathrm{GL}(n, \mathbf{F}_{p_j}) \right|.$$

The associated residual system  $\mathcal{F}^* = \{\Gamma \cap \ker r_{j,k_j}\}$  remains a normal residual system, and so we are reduced to finding a sequence of powers  $\{k_j\}$  such that

$$\lim_{j \rightarrow \infty} \frac{[\Gamma : \Gamma \cap \ker r_{j+1,k_{j+1}}]}{[\Gamma : \Gamma \cap \ker r_{j,k_j}]^{1+\delta}} = \lim_{j \rightarrow \infty} \frac{O_{j+1} p_{j+1}^{\ell_{j+1,k_{j+1}}}}{O_j^{1+\delta} p_j^{(1+\delta)\ell_{j,k_j}}} < 1. \quad (11)$$

We also require

$$[\Gamma : \Delta_{j,k_j}] \leq [\Gamma : \Delta_{j+1,k_{j+1}}]$$

for large values of  $j$ . This yields our second desired inequality

$$\frac{O_j p_j^{\ell_{j,k_j}}}{O_{j+1} p_{j+1}^{\ell_{j+1,k_{j+1}}}} < 1 \quad (12)$$

In order to achieve these inequalities, we select  $N > (n^2)!$  and  $C > 4$ . In addition, we may assume that  $\delta < 1/2$ . We also select  $0 < \varepsilon < \delta$ . In addition, let  $j_{d,\varepsilon}$  be such that for all  $j > j_{d,\varepsilon}$ , we have  $d < p_j^\varepsilon$ .

For a fixed  $j$ , the sequence  $\{\ell_{j,k}\}$  is non-decreasing and unbounded. That this sequence is non-decreasing follows from the short exact sequence. That this sequence is unbounded follows from

$$\bigcap_{k=1}^{\infty} \ker r_{j,k} = 1.$$

Setting  $\Delta_{j,k} = \ker r_{j,k}$ , we seek a sequence  $\{k_j\}$  such that

$$[\Gamma : \Delta_{j,k_j}] < [\Gamma : \Delta_{j+1,k_{j+1}}] < [\Gamma : \Delta_{j,k}]^{1+\delta}.$$

To achieve these inequalities, we construct the sequence  $k_j$  iteratively. To begin, select  $k_1$  so that  $\ell_{1,k_1} > N + Cn^2$ . Next, we select  $k_2$  such that

$$\ell_{1,k_1} + Cn^2 < \ell_{2,k_2} \leq \ell_{1,k_1} + (C+1)n^2.$$

As the sequence  $\{\ell_{2,k}\}$  is non-decreasing and unbounded, there is a largest integer  $i_2 > 1$  such that

$$\ell_{2,i_2} \leq \ell_{1,k_1} + Cn^2.$$

In particular, we have

$$\ell_{1,k_1} + Cn^2 < \ell_{2,i_2+1}.$$

Moreover, we have for all  $j, k$  that

$$\ell_{j,k+1} \leq \ell_{j,k} + n^2$$

by the short exact sequence. In particular, we see that

$$\ell_{1,k_1} + Cn^2 < \ell_{2,i_2+1} \leq \ell_{2,i_2} + n^2 \leq \ell_{1,k_1} + (C+1)n^2.$$

Setting  $k_2 = i_2 + 1$ , we also have

$$\ell_{2,k_2} > \ell_{1,k_1} + Cn^2 > N + 2Cn^2.$$

Continuing iteratively, we produce a sequence  $\{k_j\}$  such that for all  $j$ , we have

$$\ell_{j,k_j} + Cn^2 < \ell_{j+1,k_{j+1}} \leq \ell_{j,k_j} + (C+1)n^2$$

and

$$\ell_{j,k_j} > N + Cjn^2.$$

We claim that this sequence achieves the above inequalities. We start with the inequality

$$[\Gamma : \Delta_{j,k_j}] < [\Gamma : \Delta_{j+1,k_{j+1}}].$$

We know that

$$[\Gamma : \Delta_{j,k_j}] = O_j p_j^{\ell_{j,k_j}} < p_j^{n^2 + \ell_{j,k_j}}.$$

We also have

$$[\Gamma : \Delta_{j+1,k_{j+1}}] = O_{j+1} p_{j+1}^{\ell_{j+1,k_{j+1}}} > p_j^{\ell_{j+1,k_{j+1}}}.$$

Therefore, it suffices to have

$$\ell_{j+1,k_{j+1}} > n^2 + \ell_{j,k_j}.$$

This inequality holds since  $C > 4$  and

$$\ell_{j,k_j} + Cn^2 < \ell_{j+1,k_{j+1}}.$$

Next, we verify

$$[\Gamma : \Delta_{j+1, k_{j+1}}] < [\Gamma : \Delta_{j, k}]^{1+\delta}.$$

Again, we have

$$\begin{aligned} [\Gamma : \Delta_{j+1, k_{j+1}}] &= O_{j+1} p_{j+1}^{\ell_{j+1, k_{j+1}}} < p_{j+1}^{n^2 + \ell_{j+1, k_{j+1}}} \\ &\leq d^{n^2 + \ell_{j+1, k_{j+1}}} p_j^{n^2 + \ell_{j+1, k_{j+1}}} \\ &< p_j^{\varepsilon(n^2 + \ell_{j+1, k_{j+1}})} p_j^{n^2 + \ell_{j+1, k_{j+1}}} \\ &= p_j^{(\varepsilon+1)(n^2 + \ell_{j+1, k_{j+1}})} \\ &< p_j^{(\varepsilon+1)(n^2 + \ell_{j, k_j} + (C+1)n^2)} \\ &= p_j^{(\varepsilon+1)(\ell_{j, k_j} + (C+2)n^2)}. \end{aligned}$$

We also have

$$[\Gamma : \Delta_{j, k_j}]^{1+\delta} = O_j^{1+\delta} p_j^{(1+\delta)\ell_{j, k_j}} \geq p_j^{(1+\delta)\ell_{j, k_j}}.$$

Thus, we require

$$(1+\delta)\ell_{j, k_j} - (1+\varepsilon)(\ell_{j, k_j} + (C+2)n^2) > 0.$$

By construction, we know that

$$\ell_{j, k_j} > N + C j n^2.$$

So we have

$$\begin{aligned} (1+\delta)\ell_{j, k_j} - (1+\varepsilon)(\ell_{j, k_j} + (C+2)n^2) &= (\delta - \varepsilon)\ell_{j, k_j} - (1+\varepsilon)(C+2)n^2 \\ &> (\delta - \varepsilon)[N + C j n^2] - (1+\varepsilon)(C+2)n^2. \end{aligned}$$

By selection,  $\delta - \varepsilon > 0$  and  $N, n, C, \varepsilon$  are all constant. Thus, there exists  $j_0 > j_{d, \varepsilon}$  such that  $p_j^\varepsilon > d$  and

$$(\delta - \varepsilon)[N + C j n^2] - (1+\varepsilon)(C+2)n^2 > 1$$

for all  $j \geq j_0$ . In total, we see that the normal residual system  $\{\Delta_{j, k_j}\}_{j \geq j_0}$  satisfies the conditions needed for the proposition.  $\square$

Without the Strong Approximation Theorem, we have very little control in the above proof. For instance, it is not clear that the sequence  $\{\ell_{j,k}\}$  is strictly increasing for  $k \geq M$ , where  $M$  is a constant that is independent of  $j$ . If the group  $\Gamma$  is a cyclic subgroup  $\langle \alpha \rangle$  of  $\mathcal{O}_K^\times$ , the group of units in  $\mathcal{O}_K$ , one would hope that for all but finitely many primes  $\mathfrak{p}$ , the order of  $\alpha$  modulo  $\mathfrak{p}^2$  is  $a_{\mathfrak{p}} p$  where  $q$  is the cardinality of the residue field  $\mathcal{O}_K/\mathfrak{p}$  and  $a_{\mathfrak{p}}$  divides  $q-1$ . To put this problem into perspective, a Wieferich prime is a prime  $p$  such that  $p^2$  divides  $2^{p-1} - 1$ . It has been conjectured that only finitely many Wieferich primes exist [27] and also that infinitely many Wieferich primes exist [20]. The above hope is analogous to the finiteness of Wieferich primes. Generalization of this problem are related to the ABC conjecture (see [24]).

For a fixed prime ideal  $\mathfrak{p}$ , after the smallest power  $k_{\mathfrak{p}}$  such that  $\mathfrak{p}^{k_{\mathfrak{p}}}$  is not in primary decomposition of the ideal  $(\alpha^{q-1} - 1)$ , we do get strict growth on the prime powers  $p$  occurring in the multiplicative order of  $\alpha$  modulo  $\mathfrak{p}^k$ . This lack of control of  $k_{\mathfrak{p}}$  makes estimates of the indices  $s_j$  quite difficult. To prove Theorem 1.4 using Theorem 1.2, the indices  $s_j$  are precisely what requires control.

**2. Proof of Theorem 1.2.** We now prove Theorem 1.2. We start with a well known proposition.

**Proposition 5.2.** *If  $\Gamma$  is a finitely generated, infinite linear group, then there exists an infinite representation of  $\Gamma$  into  $\mathrm{GL}(n, K)$  for some  $n$  and finite extension  $K/\mathbf{Q}$ .*

*Proof.* If  $\Gamma$  is virtually solvable, the conclusion of Proposition 5.2 follows from [2, p.137]. Otherwise, the conclusion follows, for instance, from [19, Proposition 16.4.13] or [5, Lemma 3.1].  $\square$

With Propositions 5.1 and 5.2, we now quickly derive Theorem 1.2.

*Proof of Theorem 1.2.* According to Proposition 5.2, there exists an infinite linear representation

$$\rho : \Gamma \longrightarrow \mathrm{GL}(n, K)$$

for some  $n$  and  $K/\mathbf{Q}$  finite. According to Proposition 5.1, for any  $\delta > 0$ , there exists a normal residual system  $\mathcal{F}_\delta^\rho = \{\Delta'_j\}$  for  $\rho(\Gamma)$  such that for each  $\Delta'_j, \Delta'_{j+1}$

$$[\rho(\Gamma) : \Delta'_j] \leq [\rho(\Gamma) : \Delta'_{j+1}] \leq [\rho(\Gamma) : \Delta'_j]^{1+\delta}.$$

Setting

$$\mathcal{F}_\delta = \{\Delta_j : \Delta_j = \rho^{-1}(\Delta'_j), \Delta_j \in \mathcal{F}_\delta^\rho\},$$

we see that

$$[\Gamma : \Delta_j] \leq [\Gamma : \Delta_{j+1}] \leq [\Gamma : \Delta_j]^{1+\delta}.$$

That  $\mathcal{F}_\delta$  is a normal family of finite index subgroups of  $\Gamma$  follows from elementary group theory.  $\square$

## 6 Proof of Theorem 1.4

We are now ready to prove Theorem 1.4. Using the proof of Proposition 5.1, the main technical point is ensuring the coefficients  $s_j = [\Lambda_{j-1} : \Lambda_j]$  are sufficiently large. We control these values via trivial estimates. The flexibility of the proof of Proposition 5.1 allows us to use this growth condition to prove Theorem 1.4 by appealing directly to Proposition 3.4.

*Proof of Theorem 1.4.* According to Proposition 5.2, there exists an infinite representation

$$\rho : \Gamma \longrightarrow \mathrm{GL}(n, K)$$

for some  $n$  and  $K/\mathbf{Q}$  finite. By Lemma 2.5, it suffices to show that  $\mathrm{Ave}^\triangleleft(\rho(\Gamma))$  is finite. For notational simplicity, set  $\Lambda = \rho(\Gamma)$ . Finally, set  $S$  to be the coefficient ring of  $\Lambda$ .

For each  $\delta$ , from the proof of Proposition 5.1, there exists a normal residual system  $\mathcal{F}_\delta$  on  $\Lambda$  given by  $\Delta_{j,k_j} = \Lambda \cap \ker r_{j,k_j}$ , where

$$r_{j,k_j} : \mathrm{GL}(n, S) \longrightarrow \mathrm{GL}(n, S/\mathfrak{p}_{S,j}^{k_j}).$$

In addition, we have

$$|r_{j,k_j}(\Lambda)| = O_j p_j^{\ell_{j,k_j}}$$

where

$$1 \leq O_j < p_j^{n^2}.$$

We also have for constants  $N > (n^2)!$  and  $C > 4$  that

$$\ell_{j,k_j} > N + C j n^2$$

and

$$\ell_{j,k_j} + C n^2 < \ell_{j+1,k_{j+1}} \leq \ell_{j,k_j} + (C+1)n^2.$$

Finally, we will assume that  $\delta < 0$ .

To prove Theorem 1.4, we first give some trivial estimates for the size of  $s_{j,k_j}$ . For each  $i < j$ , we claim that the largest power of  $p_j$  that divides  $[\Lambda : \Delta_{i,k_i}]$  is  $p_j^{n^2}$ . To see this claim, note that if  $p_j^m$  divides  $O_i p_i^{\ell_{i,k_i}}$ , since  $p_i, p_j$  are distinct primes,  $p_j^m$  must divide  $O_i$ . However,  $O_i < p_i^{n^2}$  and  $p_i < p_j$ . Thus the claim follows. This claim thus shows that

$$[\Delta_{i,k_i} : \Delta_{i,k_i} \cap \Delta_{j,k_j}] \geq p_j^{\ell_{j,k_j} - n^2}.$$

Taking this fact over all  $i < j$ , we see that

$$[\Lambda_{j-1} : \Lambda_{j-1} \cap \Delta_{j,k_j}] \geq p_j^{\ell_{j,k_j} - (j-1)n^2}.$$

As the left hand side is  $s_j$ , we see that

$$s_{j,k_j} \geq p_j^{\ell_{j,k_j} - (j-1)n^2}.$$

We are now ready to show that  $\text{Ave}^\triangleleft(\Lambda)$  is finite. For this finiteness, by Proposition 3.4, we must show directly

$$\lim_{j \rightarrow \infty} \alpha_{j,k_j}(\mathcal{F}_\delta) = \lim_{j \rightarrow \infty} \frac{r_{j+1,k_{j+1}} s_{j+1,k_{j+1}} - r_{j+1,k_{j+1}}}{r_{j,k_j} s_{j,k_j} (s_{j,k_j} - 1)} = 0.$$

Thus, using Theorem 1.2 and the definitions of  $r_{j,k_j}, s_{j,k_j}$  (see 4), we see that

$$\begin{aligned} \frac{r_{j+1,k_{j+1}} s_{j+1,k_{j+1}} - r_{j+1,k_{j+1}}}{r_{j,k_j} s_{j,k_j} (s_{j,k_j} - 1)} &< \frac{[\Lambda : \Delta_{j+1,k_{j+1}}]}{[\Lambda : \Delta_{j,k_j}] (s_{j,k_j} - 1)} \\ &\leq \frac{[\Lambda : \Delta_{j,k_j}]^\delta}{s_{j,k_j} - 1}. \end{aligned}$$

As the term  $s_{j,k_j}$  is increasing, we are reduced to showing

$$\lim_{j \rightarrow \infty} \frac{[\Lambda : \Delta_{j,k_j}]^\delta}{s_{j,k_j}} = 0.$$

This limit is dealt with as follows.

$$\begin{aligned} \lim_{j \rightarrow \infty} \frac{[\Lambda : \Delta_{j,k_j}]^\delta}{s_{j,k_j}} &= \lim_{j \rightarrow \infty} \frac{O_j^\delta p_j^{\delta \ell_{j,k_j}}}{s_{j,k_j}} \\ &< \lim_{j \rightarrow \infty} \frac{p_j^{\delta n^2 + \delta \ell_{j,k_j}}}{p_j^{\ell_{j,k_j} - (j-1)n^2}}. \end{aligned}$$

Therefore, it suffices to have

$$\ell_{j,k_j} - (j-1)n^2 - \delta n^2 - \delta \ell_{j,k_j} > 1.$$

To that end, we have

$$\begin{aligned} \ell_{j,k_j} - (j-1)n^2 - \delta n^2 - \delta \ell_{j,k_j} &= (1-\delta)\ell_{j,k_j} - [j+\delta-1]n^2 \\ &> \frac{1}{2}\ell_{j,k_j} - jn^2 \\ &> \frac{1}{2}[N + Cjn^2] - jn^2 \\ &= \frac{1}{2}N + \left(\frac{Cj}{2} - j\right)n^2 \\ &> \frac{1}{2}N + jn^2 > n > 1. \end{aligned}$$

Hence, by Proposition 3.4,  $\text{Ave}_{\mathcal{F}_\delta}(\Lambda)$  is finite. Thus, by Lemma 2.4,  $\text{Ave}^\triangleleft(\Lambda)$  is finite and so by Lemma 2.5,  $\text{Ave}^\triangleleft(\Gamma)$  is finite.  $\square$

We obtain a different proof of Theorem 1.3 as  $\text{Ave}(\Gamma) \leq \text{Ave}^\triangleleft(\Gamma)$ .

## 7 A pair of examples

In this short section, we show that neither the finite generation nor the linearity in Theorem 1.3 can be dropped. Recall that the first Grigorchuk group  $\Gamma$  is a subgroup of automorphisms of a rooted binary tree (see [9, Ch. VIII]) and is known to be finitely generated and residual finite but not linear. The groups  $\text{SL}(n, \mathbf{Z}_p)$ , are linear and residually finite but not finitely generated for all  $p$  and  $n > 1$ .

**Theorem 7.1.** *The first Grigorchuk group  $\Gamma$  has infinite  $\text{Ave}^\triangleleft$ . The group  $\text{SL}(n, \mathbf{Z}_p)$  has infinite  $\text{Ave}^\triangleleft$  for all  $p$  and  $n > 1$ .*

*Proof.* For a normal, finite index subgroup  $\Delta$  of  $\Gamma$ , the level of  $\Delta$  is the largest  $n$  such that  $\Delta$  acts trivially on the  $n$ th level rooted binary tree. It is known that if  $\Delta$  is level  $n$ , then  $\Delta$  contains the kernel of the action on the  $(n+3)$ -level rooted binary tree (see [7]). This containment allows us to bound  $D_\Gamma^\triangleleft$  from below by a function  $D_1$  which has infinite  $L^1$ -norm. Specifically, set  $\Delta_j$  to be the kernel of the induced map on the level  $j$  rooted binary tree. We define

$$D_1(\gamma) = \begin{cases} 1, & \gamma \in \Delta_3^c \\ [\Gamma : \Delta_{j-3}], & \gamma \in \Delta_j^c \setminus \Delta_{j-1}^c, \quad j > 3. \end{cases}$$

Note that the function  $D_1$  is simply underestimating the value of  $D_\Gamma^\triangleleft$ . In particular, we see that  $D_1 \leq D_\Gamma^\triangleleft$ . Moreover, we have by construction

$$\int_{\widehat{\Gamma}} D_1 d\mu = \left( \frac{[\Gamma : \Delta_3] - 1}{[\Gamma : \Delta_3]} \right) + \sum_{j=1}^{\infty} \frac{[\Gamma : \Delta_j]}{[\Gamma : \Delta_{j+2}]} \left( \frac{[\Delta_{j+2} : \Delta_{j+3}] - 1}{[\Delta_{j+2} : \Delta_{j+3}]} \right).$$

Since the terms in latter series do not tend to zero, the above series diverges. As

$$\int_{\widehat{\Gamma}} D_1 d\mu \leq \text{Ave}^\triangleleft(\Gamma),$$

we see that  $\text{Ave}^\triangleleft(\Gamma)$  is infinite.

For  $\text{SL}(n, \mathbf{Z}_p)$ , the normal, finite index subgroups are all of the form  $\ker r_j$  where  $r_j$  is reduction modulo the  $j$ th power  $\pi^j \mathbf{Z}_p$  of the uniformizer ideal  $\pi \mathbf{Z}_p$ . In particular, these subgroups are nested and hence by Lemma 3.2,  $\text{Ave}^\triangleleft(\text{SL}(n, \mathbf{Z}_p))$  is infinite.  $\square$

Using the linear representations of  $\text{Out}(F_n)$  and  $\text{Mod}(S_g)$  obtained from acting on the first cohomology groups  $H^1(F_n, \mathbf{Z})$  and  $H^1(\pi_1(S_g), \mathbf{Z})$  respectively, we obtain from Lemma 2.5 and Proposition 5.1, that  $\text{Ave}(\text{Out}(F_n))$  and  $\text{Ave}(\text{Mod}(S_g))$  are finite. Since for  $n \geq 4$ ,  $\text{Out}(F_n)$  is not linear (see [11]), this provides an example of a non-linear, finitely generated, residually finite group with finite average. This shows that the finiteness of  $\text{Ave}(\Gamma)$  is not equivalent to the linearity of  $\Gamma$ . Finite generation is also not necessary, as  $\text{SL}(n, \widehat{\mathbf{Z}})$  has finite average but is not finitely generated. Indeed, for any residually finite, finitely generated group  $\Gamma$  with  $\text{Ave}(\Gamma)$  finite, we see that  $\text{Ave}(\widehat{\Gamma})$  is finite.

## 8 Averaging over other densities

Throughout this article we study the average defined by integrating against the Haar measure in the profinite completion. In this section we average over densities, which are not measures in general. That is, for a finitely generated group  $\Gamma$  and some density  $\delta$ , we define the *average with respect to  $\delta$*  to be

$$\text{Ave}_\delta^\triangleleft(\Gamma) = \sum_{i=1}^{\infty} i\delta((D_\Gamma^\triangleleft)^{-1}(i)).$$

For instance, given a group  $\Gamma$  generated by a finite set  $X$ , the *asymptotic density* of a subset  $S$  of  $\Gamma$  is defined to be

$$\rho(S) = \limsup_{n \rightarrow \infty} \frac{|S \cap B_{X, \Gamma}(n)|}{|B_{\Gamma, X}(n)|}.$$

This density is never additive (see [6], Example 2.2). There are other interesting notions of density, such as annular density, spherical density, and exponential density (see [6] and [16]). However when the asymptotic density exists as a limit, one can draw direct relationships between these other densities and the asymptotic density. In light of this fact, we focus on the asymptotic density.

The asymptotic density is well-behaved when we restrict attention to  $\Gamma$  satisfying:

$$\lim_{n \rightarrow \infty} \frac{|B_{\Gamma, X}(n+1)|}{|B_{\Gamma, X}(n)|} = 1 \quad (13)$$

for some finite generating set  $X$ . In this case, by [6, Section 2] the asymptotic density is both left and right invariant and we can replace the  $\limsup$  with a limit in the definition of  $\rho$  for finite index subgroups. In particular, for finite index subgroups, we have

$$\rho(\Delta) = \frac{1}{[\Gamma : \Delta]}.$$

In tandem these aforementioned facts spawn our final result.

**Theorem 8.1.** *Let  $\Gamma$  be a residually finite group with finite generating set  $X$  satisfying (13). Then for all  $n$  in  $\mathbf{N}$ ,*

$$\mu\left((D_{\hat{\Gamma}}^{\triangleleft})^{-1}(n)\right) = \rho((D_{\Gamma}^{\triangleleft}(n))^{-1}).$$

*In particular,*

$$\text{Ave}^{\triangleleft}(\Gamma) = \text{Ave}_{\rho}^{\triangleleft}(\Gamma).$$

*Proof.* For the residual system  $\text{NFI}(\Gamma)$  with the subgroups  $\Delta_j$  and  $\Lambda_j$  given by (1), for any  $i$  in  $\mathbf{N}$ , there exists  $j, k$  in  $\mathbf{N}$  such that  $j \leq k$  and

$$(D_{\Gamma}^{\triangleleft})^{-1}(i) = \Lambda_j \setminus \Lambda_k.$$

By left invariance of  $\rho$ , we have

$$\rho(\Lambda_j \setminus \Lambda_k) = \rho(\Lambda_j) - \rho(\Lambda_k).$$

Hence, we have

$$\rho(\Lambda_j \setminus \Lambda_k) = \mu(\overline{\Lambda_j} \setminus \overline{\Lambda_k}),$$

as desired.  $\square$

## 9 Zeta functions associated to finite quotients

For an infinite group  $\Gamma$ , Larsen [18] studied the zeta function

$$\zeta_\Gamma(s) = \sum_{i \in \mathcal{I}(\Gamma)} \frac{1}{i^s},$$

where

$$\mathcal{I}(\Gamma) = \{i \in \mathbf{N} : \text{NFI}_n(\Gamma) \neq \emptyset\}.$$

He related the radius of convergence to  $\min_\rho \dim \overline{\rho(\Gamma)}$ , where

$$\rho : \Gamma \longrightarrow \text{GL}(n, \mathbf{C})$$

is a linear representation with infinite image and  $\overline{\rho(\Gamma)}$  is the Zariski closure of  $\rho(\Gamma)$ . Visibly, convergence of  $\zeta_\Gamma$  is related to gaps between successive indices and our work here provides some weak results on the radius of convergence. Specifically, we can give constants  $s_\Gamma$  that ensure  $\zeta_\Gamma(s)$  diverges provided  $s \leq s_\Gamma$ .

## References

- [1] H. Bass, *Algebraic K-theory*, Springer-Verlag, 1968.
- [2] A. Borel, *Linear algebraic groups*, Springer-Verlag, 1991.
- [3] K. Bou-Rabee, *Quantifying Residual Finiteness*, to appear in J. Algebra.
- [4] K. Bou-Rabee and D. B. McReynolds, *Asymptotic growth in groups and generalized least common multiples*, submitted.
- [5] E. Breuillard and T. Gelander, *Uniform independence in linear groups*, Invent. Math. **173** (2008), 225–263.
- [6] J. Burillo and E. Ventura, *Counting Primitive Elements in Free Groups*, Geom. Dedicata **93** (2002), 132–162.
- [7] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli, *The top of the lattice of normal subgroups of the Grigorchuk group*, J. Algebra **246** (2001), 292–310.
- [8] J. S. Chahal, *Solution of the congruence subgroup problem for solvable algebraic groups*, Nagoya Math. J. **79** (1980), 141–144.

- [9] P. de la Harpe, *Topics in geometric group theory*, Chicago Lectures in Mathematics, 2000.
- [10] J. D. Dixon, M. P. F. du Sautoy, A. Mann, and D. Segal, *Analytic pro- $p$  Groups*, Cambridge Studies in Advanced Maths. **61**, 1999.
- [11] E. Formanek and C. Procesi, *The automorphism group of a free group is not linear*, J. Algebra. **149** (1992), 494–499.
- [12] A. Haar, *Der Massbegriff in der Theorie der kontinuierlichen Gruppen*, Ann. of Math. **34** (1933).
- [13] U. Hadad, *On the shortest identity in finite simple groups of Lie type*, preprint.
- [14] N. R. Howes, *Modern analysis and topology*, Spring-Verlag, 2005.
- [15] H. Iwaniec, *Almost-primes represented by quadratic polynomials*, Invent. Math. **47** (1978), 178–188.
- [16] I. Kapovich, I. Rivin, P. Schupp, and V. Shpilrain, *Densities in free groups and  $\mathbf{Z}^k$ , visible points and test elements*, Math. Res. Lett. **14** (2007), 263–284.
- [17] M. Kassabov and N. Nikolov, *Cartesian products as profinite completions*, Int. Math. Res. Not. **23** (2006), 17 pages.
- [18] M. Larsen, *How often is  $84(g-1)$  achieved?* Israel J. Math. **126** (2001), 1–16.
- [19] A. Lubotzky and D. Segal, *Subgroup growth*, Birkhäuser-Verlag, 2003 (see also on-line version available at Lubotzky's website).
- [20] D. Mirimanoff, *Sur le dernier théorème de Fermat*, J. für Math. **139** (1911), 309–324.
- [21] M. V. Nori, *On subgroups of  $\mathrm{GL}_n(\mathbf{F}_p)$* , Invent. Math. **88** (1987), 257–275.
- [22] R. Pink, *Strong approximation for Zariski dense subgroups over arbitrary global fields*, Comment. Math. Helv. **75** (2000), 608–643.
- [23] I. Rivin, *Geodesics with one self-intersection, and other stories*, preprint.

- [24] J. H. Silverman, *Wieferich's criterion and the abc-conjecture*, J. Number Theory **30** (1988), 226–237.
- [25] R. Steinberg, *Lectures on Chevalley groups*, Yale 1967.
- [26] P. Tchebychef, *Œuvres de P.L. Tchebychef*, St. Pétersbourg, Commission-aires de l'Académie impériale des sciences 1899.
- [27] A. Wieferich, *Zum letzten Fermat'schen Theorem*, J. für Math. **136** (1909), 293–302.
- [28] B. Weisfeiler, *Strong approximation for Zariski-dense subgroups of semisimple algebraic groups*, Ann. of Math. (2) **120** (1984), 271–315.
- [29] J. S. Wilson, *Profinite groups*, Oxford Press, 1997.

Department of Mathematics  
University of Chicago  
Chicago, IL 60637, USA  
email: khalid@math.uchicago.edu, dmcreyn@math.uchicago.edu